

## SEQUENCE LISTING

<110> Japan Tobacco, Inc.

<120> HCV Polymerase Suitable For Crystal Structure Analysis  
And Method For Using The Enzyme

<130> J00-0086

<140>

<141>

<150> JP P1999-188630

<151> 1999-07-02

<150> JP P1999-192488

<151> 1999-07-07

<160> 24

<170> PatentIn Ver. 2.1

- <210> 1

<211> 591

<212> PRT

<213> Hepatitis C virus

<400> 1

Ser Met Ser Tyr Thr Trp Thr Gly Ala Leu Ile Thr Pro Cys Ala Ala

1

5

10

15

Glu Glu Ser Lys Leu Pro Ile Asn Ala Leu Ser Asn Ser Leu Leu Arg

20

25

30

His His Asn Met Val Tyr Ala Thr Thr Ser Arg Ser Ala Gly Leu Arg

35

40

45

Gln Lys Lys Val Thr Phe Asp Arg Leu Gln Val Leu Asp Asp His Tyr

50

55

60

Arg Asp Val Leu Lys Glu Met Lys Ala Lys Ala Ser Thr Val Lys Ala

65

70

75

80

Lys Leu Leu Ser Val Glu Glu Ala Cys Lys Leu Thr Pro Pro His Ser

85

90

95

Ala Lys Ser Lys Phe Gly Tyr Gly Ala Lys Asp Val Arg Asn Leu Ser

100

105

110

Ser Lys Ala Val Asn His Ile His Ser Val Trp Lys Asp Leu Leu Glu

115

120

125

Asp Thr Val Thr Pro Ile Asp Thr Thr Ile Met Ala Lys Asn Glu Val

130

135

140

Phe Cys Val Gln Pro Glu Lys Gly Gly Arg Lys Pro Ala Arg Leu Ile

145

150

155

160

Val Phe Pro Asp Leu Gly Val Arg Val Cys Glu Lys Met Ala Leu Tyr

165

170

175

Asp Val Val Ser Thr Leu Pro Gln Val Val Met Gly Ser Ser Tyr Gly

180

185

190

Phe Gln Tyr Ser Pro Gly Gln Arg Val Glu Phe Leu Val Asn Thr Trp

195

200

205

Lys Ser Lys Lys Asn Pro Met Gly Phe Ser Tyr Asp Thr Arg Cys Phe

210

215

220

Asp Ser Thr Val Thr Glu Asn Asp Ile Arg Val Glu Glu Ser Ile Tyr

225

230

235

240

Gln Cys Cys Asp Leu Ala Pro Glu Ala Arg Gln Ala Ile Lys Ser Leu

245

250

255

Thr Glu Arg Leu Tyr Ile Gly Gly Pro Leu Thr Asn Ser Lys Gly Gln

260

265

270

Asn Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu Thr Thr Ser

275

280

285

Cys Gly Asn Thr Leu Thr Cys Tyr Leu Lys Ala Ser Ala Ala Cys Arg

290

295

300

Ala Ala Lys Leu Gln Asp Cys Thr Met Leu Val Asn Gly Asp Asp Leu

305

310

315

320

Val Val Ile Cys Glu Ser Ala Gly Thr Gln Glu Asp Ala Ala Ser Leu  
325                    330                    335

Arg Val Phe Thr Glu Ala Met Thr Arg Tyr Ser Ala Pro Pro Gly Asp  
340                    345                    350

Pro Pro Gln Pro Glu Tyr Asp Leu Glu Leu Ile Thr Ser Cys Ser Ser  
355                    360                    365

Asn Val Ser Val Ala His Asp Ala Ser Gly Lys Arg Val Tyr Tyr Leu  
370                    375                    380

Thr Arg Asp Pro Thr Thr Pro Leu Ala Arg Ala Ala Trp Glu Thr Ala  
385                    390                    395                    400

Arg His Thr Pro Val Asn Ser Trp Leu Gly Asn Ile Ile Met Tyr Ala  
405                    410                    415

Pro Thr Leu Trp Ala Arg Met Ile Leu Met Thr His Phe Phe Ser Ile  
420                    425                    430

- Leu Leu Ala Gln Glu Gln Leu Glu Lys Ala Leu Asp Cys Gln Ile Tyr  
435                    440                    445

Gly Ala Cys Tyr Ser Ile Glu Pro Leu Asp Leu Pro Gln Ile Ile Glu  
450                    455                    460

Arg Leu His Gly Leu Ser Ala Phe Ser Leu His Ser Tyr Ser Pro Gly

465                    470                    475                    480

Glu Ile Asn Arg Val Ala Ser Cys Leu Arg Lys Leu Gly Val Pro Pro

485                    490                    495

Leu Arg Val Trp Arg His Arg Ala Arg Ser Val Arg Ala Arg Leu Leu

500                    505                    510

Ser Gln Gly Gly Arg Ala Ala Thr Cys Gly Lys Tyr Leu Phe Asn Trp

515                    520                    525

Ala Val Lys Thr Lys Leu Lys Leu Thr Pro Ile Pro Ala Ala Ser Gln

530                    535                    540

Leu Asp Leu Ser Gly Trp Phe Val Ala Gly Tyr Ser Gly Gly Asp Ile

545                    550                    555                    560

Tyr His Ser Leu Ser Arg Ala Arg Pro Arg Trp Phe Met Leu Cys Leu

565                    570                    575

Leu Leu Leu Ser Val Gly Val Gly Ile Tyr Leu Leu Pro Asn Arg

580                    585                    590

<210> 2

<211> 1743

<212> DNA

<213> Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence:DNA encoding  
 fusion protein consisting of a portion of HCV  
 polymerase and histidine tag at the C-terminus

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(1743)

&lt;400&gt; 2

atg tca atg tcc tac aca tgg aca ggc gcc ttg atc acg cca tgc gct	48
Met Ser Met Ser Tyr Thr Trp Thr Gly Ala Leu Ile Thr Pro Cys Ala	
1	5
	10
	15

gcg gag gaa agc aag ctg ccc atc aac gcg ttg agc aac tct ttg ctg	96
Ala Glu Glu Ser Lys Leu Pro Ile Asn Ala Leu Ser Asn Ser Leu Leu	
20	25
	30

cgc cac cat aac atg gtt tat gcc aca aca tct cgc agc gca ggc ctg	144
Arg His His Asn Met Val Tyr Ala Thr Thr Ser Arg Ser Ala Gly Leu	
35	40
	45

cg <del>g</del> cag aag gtc acc ttt gac aga ctg caa gtc ctg gac gac cac	192
Arg Gln Lys Lys Val Thr Phe Asp Arg Leu Gln Val Leu Asp Asp His	
50	55
	60

tac cgg gac gtg ctc aag gag atg aag gcg aag gcg tcc aca gtt aag	240
Tyr Arg Asp Val Leu Lys Glu Met Lys Ala Lys Ala Ser Thr Val Lys	
65	70
	75
	80

gct aaa ctc cta tcc gta gag gaa gcc tgc aag ctg acg ccc cca cat 288  
 Ala Lys Leu Leu Ser Val Glu Glu Ala Cys Lys Leu Thr Pro Pro His

85 90 95

tcg gcc aaa tcc aag ttt ggc tat ggg gca aag gac gtc cggtt aac cta 336  
 Ser Ala Lys Ser Lys Phe Gly Tyr Gly Ala Lys Asp Val Arg Asn Leu

100 105 110

tcc agc aag gcc gtt aac cac atc cac tcc gtg tgg aag gac ttg ctg 384  
 Ser Ser Lys Ala Val Asn His Ile His Ser Val Trp Lys Asp Leu Leu

115 120 125

gaa gac act gtg aca cca att gac acc acc atc atg gca aaa aat gag 432  
 Glu Asp Thr Val Thr Pro Ile Asp Thr Thr Ile Met Ala Lys Asn Glu

130 135 140

gtt ttc tgt gtc caa cca gag aaa gga ggc cgt aag cca gcc cgc ctt 480  
 Val Phe Cys Val Gln Pro Glu Lys Gly Gly Arg Lys Pro Ala Arg Leu

145 150 155 160

- atc gta ttc cca gat ctg gga gtc cgt gta tgc gag aag atg gcc ctc 528  
 Ile Val Phe Pro Asp Leu Gly Val Arg Val Cys Glu Lys Met Ala Leu

165 170 175

tat gat gtg gtc tcc acc ctt cct cag gtc gtg atg ggc tcc tca tac 576  
 Tyr Asp Val Val Ser Thr Leu Pro Gln Val Val Met Gly Ser Ser Tyr

180 185 190

gga ttc cag tac tct cct ggg cag cga gtc gag ttc ctg gtg aat acc 624  
Gly Phe Gln Tyr Ser Pro Gly Gln Arg Val Glu Phe Leu Val Asn Thr

195 200 205

tgg aaa tca aag aaa aac ccc atg ggc ttt tca tat gac act cgc tgt 672  
Trp Lys Ser Lys Lys Asn Pro Met Gly Phe Ser Tyr Asp Thr Arg Cys  
210 215 220

ttc gac tca acg gtc acc gag aac gac atc cgt gtt gag gag tca att 720  
Phe Asp Ser Thr Val Thr Glu Asn Asp Ile Arg Val Glu Glu Ser Ile  
225 230 235 240

tac caa tgt tgt gac ttg gcc ccc gaa gcc aga cag gcc ata aaa tcg 768  
Tyr Gln Cys Cys Asp Leu Ala Pro Glu Ala Arg Gln Ala Ile Lys Ser  
245 250 255

ctc aca gag cgg ctt tat atc ggg ggt cct ctg act aat tca aaa ggg 816  
Leu Thr Glu Arg Leu Tyr Ile Gly Gly Pro Leu Thr Asn Ser Lys Gly  
260 265 270

cag aac tgc ggt tat cgc cgg tgc cgc gcg agc ggc gtg ctg acg act 864  
Gln Asn Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu Thr Thr  
275 280 285

agc tgc ggt aac acc ctc aca tgt tac ttg aag gcc tct gca gcc tgt 912  
Ser Cys Gly Asn Thr Leu Thr Cys Tyr Leu Lys Ala Ser Ala Ala Cys  
290 295 300

cga gct gcg aag ctc cag gac tgc acg atg ctc gtg aac gga gac gac 960

Arg Ala Ala Lys Leu Gln Asp Cys Thr Met Leu Val Asn Gly Asp Asp				
305	310	315	320	
				1008
ctc gtc gtt atc tgt gaa agc gcg gga acc caa gag gac gcg gcg agc				
Leu Val Val Ile Cys Glu Ser Ala Gly Thr Gln Glu Asp Ala Ala Ser				
325	330	335		
				1056
cta cga gtc ttc acg gag gct atg act agg tac tcc gcc ccc ccc ggg				
Leu Arg Val Phe Thr Glu Ala Met Thr Arg Tyr Ser Ala Pro Pro Gly				
340	345	350		
				1104
gac ccg ccc caa cca gaa tac gac ttg gag ctg ata aca tca tgt tcc				
Asp Pro Pro Gln Pro Glu Tyr Asp Leu Glu Leu Ile Thr Ser Cys Ser				
355	360	365		
				1152
tcc aat gtg tcg gtc gcc cac gat gca tca ggc aaa agg gtg tac tac				
Ser Asn Val Ser Val Ala His Asp Ala Ser Gly Lys Arg Val Tyr Tyr				
370	375	380		
				1200
ctc acc cgt gat ccc acc acc ccc ctc gca cgg gct gcg tgg gag aca				
Leu Thr Arg Asp Pro Thr Thr Pro Leu Ala Arg Ala Ala Trp Glu Thr				
385	390	395	400	
				1248
gct aga cac act cca gtt aac tcc tgg cta ggc aac att att atg tat				
Ala Arg His Thr Pro Val Asn Ser Trp Leu Gly Asn Ile Ile Met Tyr				
405	410	415		
				1296
gcg ccc act ttg tgg gca agg atg att ctg atg act cac ttc ttc tcc				
Ala Pro Thr Leu Trp Ala Arg Met Ile Leu Met Thr His Phe Phe Ser				

420                  425                  430

atc ctt cta gcg cag gag caa ctt gaa aaa gcc ctg gac tgc cag atc 1344  
Ile Leu Leu Ala Gln Glu Gln Leu Glu Lys Ala Leu Asp Cys Gln Ile

tac ggg gcc tgt tac tcc att gag cca ctt gac cta cct cag atc att 1392  
Tyr Gly Ala Cys Tyr Ser Ile Glu Pro Leu Asp Leu Pro Gln Ile Ile

450                    455                    460

gaa cga ctc cat ggc ctt agc gca ttt tca ctc cat agt tac tct cca 1440  
Glu Arg Leu His Gly Leu Ser Ala Phe Ser Leu His Ser Tyr Ser Pro

465                  470                  475                  480

ggg gag atc aat agg gtg gct tca tgc ctc agg aaa ctt ggg gta cca 1488  
Gly Glu Ile Asn Arg Val Ala Ser Cys Leu Arg Lys Leu Gly Val Pro

ccc ttg cga gtc tgg aga cat cg<sup>g</sup> gcc agg agc gtc cg<sup>c</sup> gct agg cta 1536  
Pro Leu Arg Val Trp Arg His Arg Ala Arg Ser Val Arg Ala Arg Leu

500                    505                    510

ctg tcc cag ggg ggg agg gcc gcc act tgt ggc aag tac ctc ttc aac 1584  
Leu Ser Gln Gly Gly Arg Ala Ala Thr Cys Gly Lys Tyr Leu Phe Asn

515                    520                    525

tgg gca gtg aag acc aaa ctc aaa ctc act cca atc ccg gct gcg tcc 1632  
Trp Ala Val Lys Thr Lys Leu Lys Leu Thr Pro Ile Pro Ala Ala Ser

530                    535                    540

cag ctg gac ttg tcc ggc tgg ttc gtt gct ggt tac agc ggg gga gac 1680  
 Gln Leu Asp Leu Ser Gly Trp Phe Val Ala Gly Tyr Ser Gly Gly Asp  
 545 550 555 560

ata tat cac agc ctg tct cgt gcc cga ccc cgc gga tcc cat cac cat 1728  
 Ile Tyr His Ser Leu Ser Arg Ala Arg Pro Arg Gly Ser His His His  
 565 570 575

cac cat cac taa taa 1743  
 His His His  
 580

<210> 3

<211> 579

<212> PRT

<213> Artificial Sequence

<223> Description of Artificial Sequence:DNA encoding  
 fusion protein consisting of a portion of HCV  
 polymerase and histidine tag at the C-terminus

<400> 3

Met Ser Met Ser Tyr Thr Trp Thr Gly Ala Leu Ile Thr Pro Cys Ala  
 1 5 10 15

Ala Glu Glu Ser Lys Leu Pro Ile Asn Ala Leu Ser Asn Ser Leu Leu  
 20 25 30

Arg His His Asn Met Val Tyr Ala Thr Thr Ser Arg Ser Ala Gly Leu  
 35 40 45

Arg Gln Lys Lys Val Thr Phe Asp Arg Leu Gln Val Leu Asp Asp His  
50 55 60  
Tyr Arg Asp Val Leu Lys Glu Met Lys Ala Lys Ala Ser Thr Val Lys  
65 70 75 80  
Ala Lys Leu Leu Ser Val Glu Glu Ala Cys Lys Leu Thr Pro Pro His  
85 90 95  
Ser Ala Lys Ser Lys Phe Gly Tyr Gly Ala Lys Asp Val Arg Asn Leu  
100 105 110  
Ser Ser Lys Ala Val Asn His Ile His Ser Val Trp Lys Asp Leu Leu  
115 120 125  
Glu Asp Thr Val Thr Pro Ile Asp Thr Thr Ile Met Ala Lys Asn Glu  
130 135 140  
Val Phe Cys Val Gln Pro Glu Lys Gly Arg Lys Pro Ala Arg Leu  
145 150 155 160  
Ile Val Phe Pro Asp Leu Gly Val Arg Val Cys Glu Lys Met Ala Leu  
165 170 175  
Tyr Asp Val Val Ser Thr Leu Pro Gln Val Val Met Gly Ser Ser Tyr  
180 185 190  
Gly Phe Gln Tyr Ser Pro Gly Gln Arg Val Glu Phe Leu Val Asn Thr  
195 200 205  
Trp Lys Ser Lys Lys Asn Pro Met Gly Phe Ser Tyr Asp Thr Arg Cys  
210 215 220  
Phe Asp Ser Thr Val Thr Glu Asn Asp Ile Arg Val Glu Glu Ser Ile  
225 230 235 240  
Tyr Gln Cys Cys Asp Leu Ala Pro Glu Ala Arg Gln Ala Ile Lys Ser  
245 250 255  
Leu Thr Glu Arg Leu Tyr Ile Gly Gly Pro Leu Thr Asn Ser Lys Gly  
260 265 270  
Gln Asn Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu Thr Thr

275	280	285
Ser Cys Gly Asn Thr Leu Thr Cys Tyr Leu Lys Ala Ser Ala Ala Cys		
290	295	300
Arg Ala Ala Lys Leu Gln Asp Cys Thr Met Leu Val Asn Gly Asp Asp		
305	310	315
Leu Val Val Ile Cys Glu Ser Ala Gly Thr Gln Glu Asp Ala Ala Ser		
325	330	335
Leu Arg Val Phe Thr Glu Ala Met Thr Arg Tyr Ser Ala Pro Pro Gly		
340	345	350
Asp Pro Pro Gln Pro Glu Tyr Asp Leu Glu Leu Ile Thr Ser Cys Ser		
355	360	365
Ser Asn Val Ser Val Ala His Asp Ala Ser Gly Lys Arg Val Tyr Tyr		
370	375	380
Leu Thr Arg Asp Pro Thr Thr Pro Leu Ala Arg Ala Ala Trp Glu Thr		
385	390	395
Ala Arg His Thr Pro Val Asn Ser Trp Leu Gly Asn Ile Ile Met Tyr		
405	410	415
Ala Pro Thr Leu Trp Ala Arg Met Ile Leu Met Thr His Phe Phe Ser		
420	425	430
Ile Leu Leu Ala Gln Glu Gln Leu Glu Lys Ala Leu Asp Cys Gln Ile		
435	440	445
Tyr Gly Ala Cys Tyr Ser Ile Glu Pro Leu Asp Leu Pro Gln Ile Ile		
450	455	460
Glu Arg Leu His Gly Leu Ser Ala Phe Ser Leu His Ser Tyr Ser Pro		
465	470	475
Gly Glu Ile Asn Arg Val Ala Ser Cys Leu Arg Lys Leu Gly Val Pro		
485	490	495
Pro Leu Arg Val Trp Arg His Arg Ala Arg Ser Val Arg Ala Arg Leu		
500	505	510

Leu Ser Gln Gly Gly Arg Ala Ala Thr Cys Gly Lys Tyr Leu Phe Asn  
515 520 525  
Trp Ala Val Lys Thr Lys Leu Lys Leu Thr Pro Ile Pro Ala Ala Ser  
530 535 540  
Gln Leu Asp Leu Ser Gly Trp Phe Val Ala Gly Tyr Ser Gly Gly Asp  
545 550 555 560  
Ile Tyr His Ser Leu Ser Arg Ala Arg Pro Arg Gly Ser His His His  
565 570 575  
His His His

<210> 4

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Artificially synthesized primer sequence, 5BNde1FW.

<220>

-> <221> primer\_bind

<222> (1)..(30)

<400> 4

catatgtcaa tgtcctacac atggacagcc

30

<210> 5  
<211> 57  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:Artificially synthesized primer sequence, 5B570HRV.

<220>  
<221> primer\_bind  
<222> (1)..(57)

<400> 5  
ttattagtga tggtgatgg t gatgggatcc gcggggtcgg gcacgagaca ggctgtg 57

<210> 6  
<211> 57  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:Artificially synthesized primer sequence, 5B552HRV.

<220>  
<221> primer\_bind  
<222> (1)..(57)

<400> 6

ttattagtga tggatggatgtt gatggatcc aacgaaccag ccggacaagt ccagctg 57

<210> 7

<211> 57

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Artificially synthesized primer sequence, 5B544HRV.

<220>

<221> primer\_bind

<222> (1)..(57)

<400> 7

ttattagtga tggatggatgtt gatggatcc ctgggacgca gccgggattt gagttag 57

<210> 8

<211> 67

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Artificially

synthesized primer sequence, 5B536HRV.

<220>

<221> primer\_bind

<222> (1)..(67)

<400> 8

ttattagtga tggtgatgg t gatggatcc gagttttagt ttggccttca ctgccagg 60

gaagagg

67

<210> 9

<211> 60

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Artificially  
synthesized primer sequence, 5B531HRV.

<220>

<221> primer\_bind

<222> (1)..(60)

<400> 9

ttattagtga tggtgatgg t gatggatcc cttcaactgcc cagttgaaga ggtacttgcc 60

<210> 10

<211> 52

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Artificially synthesized primer sequence, 5B591HRV.

<220>

<221> primer\_bind

<222> (1)..(52)

<400> 10

ttattaaatgg tcatggtgat ggtgtccgga tcgattgggg agcaggtaga tg

52

<210> 11

<211> 8

<212> PRT

<213> Hepatitis C virus

<400> 11

Xaa Xaa Xaa Leu Xaa Xaa Trp Phe Xaa

1

5

<210> 12

<211> 8

<212> PRT

<213> Hepatitis C virus

<400> 12

Lys Asp Leu Ser Gly Trp Phe Lys

1 5

<210> 13

<211> 9

<212> PRT

<213> Hepatitis C virus

<400> 13

Lys Lys Asp Leu Ser Gly Trp Phe Lys

1 5

<210> 14

<211> 8

<212> PRT

<213> Hepatitis C virus

<400> 14

Lys Asp Leu Ser Gly Trp Phe Val

1 5

<210> 15

<211> 8

<212> PRT

<213> Hepatitis C virus

<400> 15

Leu Asp Leu Ser Gly Trp Phe Lys

1 5

<210> 16

<211> 8

<212> PRT

<213> Hepatitis C virus

<400> 16

Leu Asp Leu Ser Gly Trp Phe Val

1 5

<210> 17

<211> 7

<212> PRT

<213> Hepatitis C virus

<400> 17

Asp Leu Ser Gly Trp Phe Val

1 5

<210> 18

<211> 6

<212> PRT

<213> Hepatitis C virus

<400> 18

Asp Leu Ser Gly Trp Phe

1 5

<210> 19

<211> 6

<212> PRT

<213> Hepatitis C virus

<400> 19

Leu Ser Gly Trp Phe Val

1 5

<210> 20

<211> 5

<212> PRT

<213> Hepatitis C virus

<400> 20

Leu Ser Gly Trp Phe

1 5

<210> 21  
<211> 6  
<212> PRT  
<213> Hepatitis C virus

<400> 21  
Leu Ser Gly Trp Phe Lys  
1 5

<210> 22  
<211> 6  
<212> PRT  
<213> Hepatitis C virus

<400> 22  
Lys Leu Ser Gly Trp Phe  
1 5

<210> 23  
<211> 5  
<212> PRT  
<213> Hepatitis C virus

<400> 23  
Leu Gly Gly Trp Phe  
1 5

<210> 24

<211> 5

<212> PRT

<213> Hepatitis C virus

<400> 24

Leu Ser Asp Trp Phe

1

5